

Łukasz Rauch
ResearcherID G-4899-2013
Scopus Author ID 22036526500
Scopus Author ID 55387245500
ORCID: <http://orcid.org/0000-0001-5366-743X>

Kraków, 12.10.2015

Załącznik nr 3
Do wniosku o przeprowadzenie postępowania habilitacyjnego
Autoreferat w języku polskim

1. Imię i Nazwisko.

ŁUKASZ RAUCH

2. Posiadane dyplomy, stopnie naukowe/ artystyczne - z podaniem nazwy, miejsca i roku ich uzyskania oraz tytułu rozprawy doktorskiej.

- 03.2007r.: dr inż. nauk technicznych w dyscyplinie informatyka, praca obroniona na Wydziale EAIiE Akademii Górniczo-Hutniczej, temat pracy: „Design of new algorithms, based on the particles dynamics, dedicated to images and multidimensional data processing”.
- 06.2003 r.: mgr inż. ekonomii, absolwent Wydziału Zarządzania Akademii Górniczo-Hutniczej, temat pracy magisterskiej: „Projektowanie informatycznych systemów wspomagania kontaktów z klientem”.
- 05.2002 r.: mgr inż. informatyki, absolwent Wydziału EAIiE Akademii Górniczo-Hutniczej, temat pracy magisterskiej: „Rozpoznawanie mowy przy użyciu metody fraktalnej”.

3. Informacje o dotychczasowym zatrudnieniu w jednostkach naukowych.

- 10.2003 – obecnie: **Miejsce zatrudnienia:** AGH Wydział Inżynierii Metali i Informatyki Przemysłowej, **Sposób zatrudnienia:** umowa o pracę na czas nieokreślony na stanowisku adiunkta, **Zakres pracy:** badania naukowe, dydaktyka (Java, Technologie Aplikacji Internetowych, Technologie informacyjne, Podstawy Informatyki, C++, Inżynieria Internetu, Rozpoznawanie Obrazów, Systemy Operacyjne, Matematyka Dyskretna), organizacja (współpraca przy projektach unijnych oraz krajowych, tworzenie propozycji projektowych, zarządzanie projektami, organizacja konferencji naukowych).
- 03.2013 – obecnie: **Miejsce zatrudnienia:** Akademickie Centrum Komputerowe Cyfronet AGH, **Sposób zatrudnienia:** umowa o pracę na czas trwania projektu, **Zakres pracy:** realizacja prac badawczych w ramach projektów PLGrid Plus (Dziedzinnieo zorientowane usługi i zasoby infrastruktury PL-Grid dla wspomagania Polskiej Nauki w Europejskiej Przestrzeni Badawczej),

PLGrid NG (Dziedzinowe Usługi Nowej Generacji w Infrastrukturze PL-Grid dla Polskiej Nauki) oraz VirtRoll (Virtual Rolling Mill).

02.2004 – 02.2011: *Miejsce zatrudnienia*: Wyższa Szkoła Zarządzania i Bankowości w Krakowie, *Sposób zatrudnienia*: umowa o dzieło, *Zakres pracy*: dydaktyka – Języki i Techniki Programowania (C++, Java),

4. Wskazanie osiągnięcia wynikającego z art. 16 ust. 2 ustawy z dnia 14 marca 2003r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (Dz. U. nr 65, poz. 595 ze zm.):

a) tytuł osiągnięcia naukowego/artystycznego,

Metodyka obliczeń wieloskalowych z wykorzystaniem heterogenicznych architektur sprzętowych.

b) (autor/autorzy, tytuł/tytuły publikacji, rok wydania, nazwa wydawnictwa, recenzenci wydawniczy),

Publikacje w recenzowanych czasopismach i w recenzowanych materiałach konferencyjnych:

- [1] **Rauch L.**, Heterogeneous hardware implementation of Molecular Static method for modeling of interatomic behaviour, *Procedia Computer Science*, 18, 2013, 1057-1067.
- [2] **Rauch L.**, Bachniak D., Modeling of the nanostructural defects by implementation of molecular static method for heterogeneous hardware architectures, *Computer Methods in Materials Science*, 13(1), 2013, 16-21.
- [3] **Rauch L.**, Madej L., Spytkowski P., Golab R., Development of the cellular automata framework dedicated for metallic materials microstructure evolution models, *Archives of Civil and Mechanical Engineering*, 15(1), 2015, 48-61, IF: 1.793
- [4] Bzowski K., **Rauch L.**, Implementation of heat transfer model for ARM architectures, *Computer Methods in Materials Science*, 15(3), 2015, in print.
- [5] **Rauch L.**, Rodzaj A., Scheduling of computations between GPGPU and CPU for cellular automata framework used in multiscale modelling, *SpringSim '12, Conference Proceedings*, online publication, 2012.
- [6] **Rauch L.**, Dynamic load balancing for fully coupled multiscale modelling, *Journal of Computational Science*, in review, IF: 1.231.
- [7] **Rauch L.**, Pernach M., Bzowski K., Pietrzyk M., On application of shape coefficients to creation of the statistically similar representative element of DP steels, *Computer Methods in Materials Science*, 11(4), 531–541, 2011.
- [8] **Rauch L.**, Pernach M., Bzowski K., Pietrzyk M., Application of the statistically similar representative volume element to description of microstructure of DP steel, *Rudy i Metale Nieżelazne*, 56(11), pp. 704–709, 2011 (in polish).
- [9] **Rauch L.**, Pietrzyk M., Improvement of efficiency of multiscale modelling by application of statistical representation of microstructure, *MEFORM'2013: simulation von umformprozessen*, pp. 157–166.
- [10] **Rauch L.**, Bzowski K., Szeliga D., Pietrzyk M., Creation of statistically similar

- representative volume element using PLGrid environment, Cracow'12 Grid Workshop, 2012, pp. 101–102.
- [11] **Rauch L.**, Szeliga D., Bachniak D., Bzowski K., Pietrzyk M., Application of sensitivity analysis to grid-based procedure dedicated to creation of SSRVE, eScience on distributed computing infrastructure: achievements of PLGrid Plus domain-specific services and tools, Lecture Notes in Computer Science, pp. 376–377, 2014.
- [12] **Rauch L.**, Szeliga D., Bachniak D., Bzowski K., Pietrzyk M., Słota R., Kitowski J., Identification of multi-inclusion Statistically Similar Representative Volume Element for advanced high strength steels by using data farming approach, Procedia Computer Science, 51, pp. 924–933, 2015.
- [13] **Rauch L.**, Bzowski K., Bachniak D., Pietrzyk M., Robust multiscale modelling of two-phase steels on heterogeneous hardware infrastructures by using statistically similar representative volume element, Archives of Metallurgy and Materials, 2015, in print, IF: 1.090
- [14] Bachniak D., **Rauch L.**, Application of Isogeometric Analysis to multiscale modelling of dual phase materials by using Statistically Similar Representative Volume Element, III international conference on Isogeometric Analysis IGA 2015, p. 104.
- [15] Bachniak D., **Rauch L.**, Application of Isogeometric Analysis to modeling of elastic deformation in dual phase materials by using Statistically Similar Representative Volume Element on heterogeneous hardware devices, Computer Methods in Materials Science, 15(3), 2015, in print.
- [16] **Rauch L.**, Bzowski K., Rodzaj A., OpenCL implementation of Cellular Automata Finite Element (CAFE) method, Parallel Processing and Applied Mathematics, Lecture Notes in Computer Science, 381-390, 2012.
- [17] **Rauch L.**, Bzowski K., Rodzaj A., Application of Heterogeneous Computing to CAFE Simulations of Production Processes, 7th IEEE International Conference on e-Science: Stockholm, 2011, 74-80.
- [18] Nowak P., **Rauch L.**, Efficient Cellular Automata model for prediction of damage of hot forging tools due to thermal fatigue, Computer Methods in Materials Science, 14(4), 2014, pp. 197-205.
- [19] Ambroziński M., Bzowski K., **Rauch L.**, Pietrzyk M., Application of statistically similar representative volume element in numerical simulations of crash box stamping, Archives of Civil and Mechanical Engineering, 12(2), 126–132, 2012, IF: 0.963.
- [20] **Rauch L.**, The role of heterogeneous hardware architectures in multiscale modeling, PLASTMET'2012: integrated studies of foundations of plastic deformation of metals, pp. 1-2, 2012.
- [21] Pietrzyk M., Ambroziński M., Bzowski K., **Rauch L.**, Kusiak J., Problem of efficiency of computer aided design of metal forming production cycle, XXXII. Verformungskundliches Kolloquium : Leoben, 2013, pp. 69-76.
- [22] **Rauch L.**, Kuziak R., Pietrzyk M., From High Accuracy to High Efficiency in Simulations of Processing of Dual-Phase Steels, Metallurgical and Material Transactions B, 45(2), 497-506, 2014. DOI: 10.1007/s11663-013-9926-5, IF: 1.212
- [23] **Rauch L.**, Liput J., Imiołek K., Król D., Słota R., Kitowski J., Numerical simulations of metal forming production processes and cycles by using heterogeneous computing infrastructures, Cracow'12 Grid Workshop, pp. 125–126, 2014.

Książka:

- [24] Pietrzyk M., Madej L., **Rauch L.**, Szeliga D., Computational Materials Engineering, Achieving High Accuracy and Efficiency in Metals Processing Simulations, Elsevier, Inc., ISBN: 978-0-12-416707-0, 2015.

c) omówienie celu naukowego ww. pracy/prac i osiągniętych wyników wraz z omówieniem ich ewentualnego wykorzystania.

Motywacja

Rozwój hybrydowych oraz hierarchicznych metod modelowania wieloskalowego, ich rosnąca złożoność obliczeniowa, a jednocześnie ogromne zapotrzebowanie na moc obliczeniową powodują, iż kluczowe stało się podjęcie badań mających na celu optymalizację takich obliczeń zarówno pod kątem rozwiązań implementacyjnych jak i sprzętowych. Obecnie obserwuje się dynamiczny rozwój sprzętu obliczeniowego polegający na tworzeniu coraz nowszych architektur z wykorzystaniem tzw. sprzętowych akceleratorów obliczeniowych. W tym trendzie wiodącą rolę odgrywają popularne obecnie procesory takie jak GPGPU, CBEA, XEON Phi lub FPGA, ale również coraz szerzej wykorzystywane do obliczeń procesory ARM i złożone jednostki typu APU będące fuzją procesorów i kart graficznych w jednym układzie scalonym. Wymienione rozwiązania zapewniają przyspieszenie obliczeń poprzez swoją specyficzną budowę składającą się m.in. z odrębnych multi-procesorów strumieniowych, które zawierają szereg procesorów skalarnych realizujących właściwe obliczenia. Urządzenia te często wchodzi w skład większych rozwiązań sprzętowych tworząc kompleksowe heterogeniczne architektury sprzętowe, złożone z wielu węzłów obliczeniowych o różnym wyposażeniu. Bogate możliwości konfiguracji takiego sprzętu, a także możliwości rozproszenia i zrównoleglenia obliczeń oraz konfiguracji metod wieloskalowych skłoniły mnie do podjęcia badań nad stworzeniem tytułowej metodyki obliczeń wieloskalowych dla heterogenicznych architektur sprzętowych. Opracowana metodyka umożliwia m.in. statyczne oraz dynamiczne przydzielanie zasobów sprzętowych dla metod, które zachowują się w różny sposób na sprzęcie o różnej budowie i charakterystyce wydajnościowej, a także zarządzanie metodami numerycznymi podczas obliczeń. Ponadto metodyka ma szeroki wpływ na różne aspekty życia poprzez poprawę efektywności wykonywania symulacji numerycznych przy zachowaniu wysokiej jakości wyników.

Pod kątem przyszłej realizacji badań wybrałem metody numeryczne wykorzystywane powszechnie w wielu dziedzinach nauki m.in. w medycynie, inżynierii biomedycznej, budownictwie, ekonomii, mechanice, inżynierii materiałowej i metalurgii. Jako przykłady zastosowania metodyki w praktyce wybrałem dwie ostatnie z wymienionych dyscyplin nauk technicznych tj. inżynierię materiałową i metalurgię, które po informatyce stanowią drugi z obszarów moich zainteresowań naukowych.

Cele naukowe

Głównym założeniem prac umożliwiających opracowanie metodyki obliczeń wieloskalowych przy użyciu heterogenicznych architektur sprzętowych było rozwiązanie problemów związanych m.in. z aspektami projektowania, implementacji, rozmieszczenia i zarządzania programami, których docelowym środowiskiem uruchomieniowym są środowiska sprzętowe składające się z węzłów o zróżnicowanej budowie. Realizacja tak postawionego celu wymagała przede wszystkim opracowania i implementacji procedur numerycznych wykorzystywanych w jednej skali, które mogłyby zostać połączone ze sobą w bardziej złożone metody wieloskalowe i poddane wnikliwej analizie wydajnościowej. Zdecydowałem się zatem na implementację kluczowych pod kątem wydajności elementów następujących metod: Statyki i Dynamiki Molekularnej (MS, ang. Molecular Static / MD, ang. Molecular Dynamic), Automatów Komórkowych (CA, ang. Cellular Automata) oraz Metody Elementów Skończonych (FEM, ang. Finite Element Method). Następnie postanowiłem wykonać analizę wydajności obliczeniowej stworzonych metod pojedynczo oraz w różnych kombinacjach ze

sobą nawzajem jak i z dostępnym sprzętem komputerowym. Stworzone metody zostały połączone w następujący sposób: MS/MD + FE do symulacji wieloskalowej nano-makro lub nano-mikro, CA +FE (podejście CAFE) do symulacji mikro-makro oraz FE+FE (podejście FE²) do symulacji mikro-makro, aby jak najlepiej odzwierciedlić ich wykorzystanie w praktyce. Ostatnim celem było opracowanie strategii, która umożliwia:

- a) optymalne wykorzystanie lub dobór sprzętu komputerowego dla wybranego zadania obliczeniowego
- b) optymalną konfigurację i zarządzanie zadaniami obliczeniowymi (np.: podział i reprezentację domeny obliczeniowej, dynamiczne równoważenie obciążenia obliczeniowego) dla dostępnego sprzętu komputerowego.

Szczegółowe cele prac, które zostały postawione i zrealizowane są następujące:

1. *Stworzenie modułu obliczeniowego do symulacji numerycznych w skali nano:*

- Zaprojektowanie i implementacja algorytmów pozwalających na składanie układu równań po stronie urządzenia obliczeniowego,
- Zaprojektowanie oraz implementacja algorytmów rozwiązujących układy równań na urządzeniach obliczeniowych,
- Zaprojektowanie efektywnych algorytmów zarządzania pamięcią dla grupy wątków pracujących wewnątrz jednego procesora strumieniowego.

2. *Stworzenie modułu obliczeniowego do symulacji numerycznych w skali mikro:*

- Adaptacja istniejącej platformy programistycznej do automatów komórkowych do języka OpenCL i heterogenicznych platform sprzętowych,
- Opracowanie algorytmów efektywnego zarządzania pamięcią dla algorytmów bazujących na metodzie automatów komórkowych,
- Opracowanie metod redukcji redundancji danych w szczególności dla wieloskalowej reprezentacji materiałów,
- Opracowanie dekompozycji problemu dla różnych węzłów obliczeniowych z wykorzystaniem podejścia Message Passing Interface (MPI),
- Opracowanie oprogramowania do symulacji modeli materiałowych w skali mikro z wykorzystaniem metody Monte Carlo dla heterogenicznych architektur sprzętowych.

3. *Stworzenie modułu obliczeniowego dla skali makro*

- Opracowanie projektu równoległego tworzenia siatek dla różnych kształtów z różnymi elementami,
- Tworzenie efektywnego algorytmu przebudowy siatki elementów, które mogą być kluczowe dla zmieniającego się kształtu symulowanego obiektu w kolejnych iteracjach obliczeń,
- Analiza możliwości redukcji złożoności domeny obliczeniowej dla FEM poprzez zastąpienie konwencjonalnego podejścia metody FEM analizą izogeometryczną (ang. Isogeometric Analysis, IGA),
- Wykorzystanie podejścia Statystycznie Podobnego Reprezentatywnego Elementu Objętościowego (ang. Statistically Similar Representative Volume Element, SSRVE) do redukcji złożoności reprezentacji domeny obliczeniowej,
- Połączenie podejścia SSRVE i IGA dla heterogenicznych platform sprzętowych,
- Efektywne tworzenie lokalnych macierzy sztywności dla heterogenicznych architektur sprzętowych,
- Opracowanie algorytmu dekompozycji problemu z wykorzystaniem podejścia Message Passing Interface (MPI).

4. *Analiza efektywności modułów obliczeniowych w symulacjach pojedynczej skali*

- Wyznaczenie reprezentatywnych algorytmów dla modułów obliczeniowych w skali nano, mikro i makro,
- Analiza skalowalności stworzonego oprogramowania oraz ocena wiarygodności stworzonych modeli przy zmieniającym się rozmiarze danych,
- Wykonanie analizy wrażliwości modeli dla różnych parametrów sprzętowych.

5. *Optymalizacja efektywności obliczeń dla podejść sprzężonych oraz pół-sprzężonych.*

- Opracowanie metodyki dla rozmieszczenia sprzężonych i pół-sprzężonych podejść do modelowania wieloskalowego CA-FE, MS-FE, MS-CA-FE,
- Opracowanie metodyki konfiguracji heterogenicznych platform sprzętowych dla obliczeń wieloskalowych.
- Opracowanie projektu oraz implementacja wybranego solwera układów równań liniowych dla układu macierzy programowalnych bramek (ang. Field Programming Gate Array) z wykorzystaniem urządzeń pracujących w ACK Cyfronet AGH,
- Wykonanie projektu oraz implementacja frontального solwera układów równań liniowych dla metody elementów skończonych oraz IGA, weryfikacja i porównanie działania solwerów dla różnych procesorów GPGPU.

6. *Zastosowania praktyczne i weryfikacja wyników.*

- Weryfikacja opracowanych modeli w skali pojedynczej, obliczeń sprzężonych oraz pół-sprzężonych modeli.
- Przeprowadzenie badania zużycia energii elektrycznej dla urządzeń w architekturze heterogenicznej. Wykonanie analizy możliwości poprawy procentowego wykorzystania mocy obliczeniowych urządzeń dla metod MS/MD, CA oraz FE.

Ważnym aspektem podjęcia badań w tym kierunku był również rozwój multi-architekturalnych technologii implementacyjnych jak np. OpenCL, dzięki którym możliwe było opracowanie uniwersalnych kodów źródłowych kompilowalnych i uruchamialnych w różnych środowiskach sprzętowych. Dlatego też większość wymienionych powyżej celów związanych z implementacją oprogramowania zrealizowana została z wykorzystaniem technologii OpenCL, aby kod źródłowy był homogeniczny przy heterogenicznym sprzęcie komputerowym.

Efekty końcowe i osiągnięcia

Opracowana metodyka prowadzenia obliczeń wieloskalowych z wykorzystaniem heterogenicznych architektur sprzętowych obejmuje następujące elementy będące głównymi efektami badań:

1. ***Wybór optymalnego rozwiązania pomiędzy doбором architektury sprzętowej a konfiguracją wieloskalowych metod obliczeniowych w podejściu półsprzężonym.***

Wieloskalowe obliczenia w podejściu półsprzężonym charakteryzują się wymianą danych w jednym kierunku (najczęściej od skali wyższej do skali niższej), dzięki czemu obliczenia mogą być wykonane sekwencyjnie lub na wół równoległe (w następstwie kolejnych iteracji). Powoduje to, iż metody obliczeniowe w wyższej skali mogą być konfigurowane i uruchamiane niezależnie od skali niższej. Dlatego też w tej części opracowania metodyki skupiłem się na analizie metod obliczeniowych w jednej skali w zależności od wybranej architektury sprzętowej. W ramach prac implementacyjnych opracowałem procedury

umożliwiający modelowanie numeryczne metodami MS, MD, CA oraz FEM. Implementacja opracowana została w technologii OpenCL, dzięki czemu możliwe było wykonanie testów efektywności tego samego oprogramowania z wykorzystaniem różnego sprzętu komputerowego. Wykonane testy pokazały, iż w zależności od metod w pojedynczej skali lub ich połączenia w rozwiązaniu wieloskalowym, pomimo homogenicznego modelu implementacyjnego algorytmy zachowują się w różny sposób ze względu na różne parametry uruchomieniowe jak np.: Local Work Size (LWS) i Global Work Size, a także mają różne przyspieszenia oraz skalowalność dla różnych urządzeń. Pomiary zostały wykonane dla wybranych procesorów CPU, kart graficznych oraz kart Intel Xeon Phi dla następujących metod:

- a. Statyki i dynamiki molekularnej – przeanalizowane zostały trzy różne podejścia do dekompozycji domeny obliczeniowej bazującej na podziale atomów, sił oraz przestrzeni. Ze względu na parametry wybranych kart graficznych, radziły one sobie znacznie szybciej z obliczeniami niż konwencjonalne procesory CPU dla tych samych rozmiarów przestrzeni. Porównane zostały zarówno czasy, jak i szczytowa efektywność obliczeniowa w jednostkach GFlops. Podobnie sytuacja wygląda w przypadku karty Intel Xeon Phi, które były znacząco szybsze od konwencjonalnych procesorów CPU. Szczegóły wyników zawierają prace [1,2].

Głównym osiągnięciem tej części prac było uzyskanie odpowiedzi na pytanie, w jaki sposób dobrać optymalnie sprzęt do wykorzystywanej algorytmiki lub jak wykonać najlepszą dekompozycję domeny obliczeniowej dla z góry zadanej konfiguracji sprzętowej. Dzięki temu uzyskane zostały bardzo dobre przyspieszenie i skalowalność oprogramowania na wielu poziomach architektury sprzętowej tj. między węzłami obliczeniowymi, między urządzeniami wewnątrz pojedynczego węzła oraz wewnątrz urządzeń obliczeniowych między procesorami strumieniowymi. Dodatkowo zbadany został również wpływ kluczowych parametrów zaimplementowanych algorytmów, jak np.: funkcje potencjałów międzycząsteczkowych, na wydajność oprogramowania. Wskazówki uzyskane dzięki tym badaniom są jednym z elementów opracowanej metodyki. Stworzone oprogramowanie wykorzystane zostało do analizy wydajności podejść wieloskalowych w pełni i pół-sprzężonych.

- b. Automatów komórkowych – opracowana została uniwersalna platforma CA umożliwiająca automatyczne generowanie kerneli obliczeniowych na podstawie modeli numerycznych zaimplementowanych z wykorzystaniem API opracowanej platformy. Dzięki temu możliwe było wykorzystanie pewnych uniwersalnych elementów automatów komórkowych jak przestrzeń automatów, reguły przejścia i stany do tworzenia różnych modeli numerycznych. Modele te następnie są konwertowane do postaci kerneli w technologii OpenCL. Platforma wspiera przede wszystkim homogeniczne automaty komórkowe, dla których opracowano kilka modeli o różnej złożoności obliczeniowej. Wykonane badania wykazały, iż podobnie jak w przypadku algorytmów skali nano, ze względu na przyspieszenie i skalowalność lepiej wypadają karty obliczeniowe niż wielordzeniowe procesory konwencjonalne. Jednakże reguła ta ma zastosowanie tylko w przypadku, kiedy przestrzeń obliczeniowa mieści się w całości w pamięci urządzenia obliczeniowego. W przeciwnym razie przesyłanie danych jest wąskim gardłem obliczeń i wówczas korzystniejszym rozwiązaniem jest wykonanie obliczeń za pomocą CPU oraz pamięci RAM. Opis platformy oraz wyniki uzyskane przez kernele generowane dla architektur wybranych kart graficznych znaleźć można w publikacji [3]. Stworzone oprogramowanie umożliwia, podobnie jak w przypadku podpunktu (a), wielopoziomowe zrównoleglenie obliczeń. Zrównoleglenie zrealizowane zostało z wykorzystaniem

dekompozycji przestrzeni automatów komórkowych dla różnych wariantów – równomierny podział przestrzeni w różnych wymiarach (podział po osiach, X, Y i Z oddzielnie jak i w różnych kombinacjach) oraz dowolny podział przestrzeni np.: w zależności od budowy wirtualnej mikrostruktury modelowanego materiału (podział po ziarnach oraz fazach materiału, gdzie obliczenia mogą być bardziej intensywne w zależności od obszaru).

Głównym osiągnięciem jest automatyzacja generowania kerneli i konfiguracji prowadzonych obliczeń dla różnych urządzeń. Dodatkowo, zastosowanie redundancji danych w dodatkowych klasach platformy obliczeniowej umożliwiło redukcję złożoności obliczeniowej. Narzut na dodatkowe dane jest niewielki w porównaniu z przyspieszeniem obliczeń, które dzięki niemu zostaje uzyskane. W ramach weryfikacji zaimplementowane zostały modele rozwoju mikrostruktury materiału metalicznego: rekrytalizacji dynamicznej oraz pęknięcia zmęczeniowego.

- c. Metody elementów skończonych – analizie poddany został proces składania lokalnych oraz globalnej macierzy MES oraz rozwiązywanie układów z wykorzystaniem solverów z metodą gradientów sprzężonych, stabilizowanych gradientów sprzężonych, GMRES oraz dekompozycji LU. Solverzy iteracyjne w odróżnieniu od poprzednio analizowanych metod wykazały bardzo dobre przyspieszenie i skalowalność dla konwencjonalnych procesorów, a zdecydowanie gorsze wyniki obserwowane były dla kart graficznych oraz urządzeń Intel Xeon Phi. Wprowadzono ulepszenia metod polegające na składaniu układów równań bezpośrednio na urządzeniach obliczeniowych, co znacząco podniosło szybkość działania oprogramowania, jednak nie na tyle, aby wyprzedzić konwencjonalne procesory CPU. Dlatego też ze względu na hierarchiczną budowę urządzeń obliczeniowych oraz pamięć mniejszą niż ogólnodostępna pamięć RAM, szczególny nacisk położony został na solverzy frontalne bazujące na gramatykach oraz eliminacji Gaussa. Otrzymane wyniki pokazują, iż pomimo bardzo dobrego zrównoleglenia, nadal przewagę w efektywności obliczeniowej mają procesory CPU nad innymi analizowanymi urządzeniami. Dodatkowym wynikiem prac było wykonanie projektu oraz implementacji solvera gradientów sprzężonych dla urządzeń FPGA. Implementację wykonano w języku *Impulse C* pod kątem reprogramowalnej platformy Picco Virtex 6. Zwieńczeniem prac było opracowanie procedur uruchomienia oprogramowania na sprzęcie. Ponadto, wykonano również analizę dla niezwykle popularnych architektur ARM, której wyniki opublikowano w pracy [4]. Analiza MES wykonana została głównie pod kątem energooszczędności rozwiązań ARM w porównaniu do innych architektur sprzętowych.

W ramach prac nad redukcją złożoności domeny obliczeniowej wykonane zostały badania nad wykorzystaniem Analizy Izometrycznej (IGA) do reprezentacji kształtu domeny oraz siatki elementów. Opracowane zostały metody do tworzenia siatek elementów skończonych oraz elementów IGA, a następnie opracowane zostały algorytmy do p -, h - i k -adaptacji. Zaimplementowane zostały również metody numeryczne do składania lokalnych i globalnego układu równań bezpośrednio na urządzeniach obliczeniowych, co umożliwiło znaczne przyspieszenie działania oprogramowania. Do rozwiązywania układów równań wykorzystana została biblioteka ViennaCL oraz opracowany został własny zrównoleglony solver frontalny, pozwalający na prowadzenie obliczeń w ramach jednego węzła obliczeniowego jak i wielu węzłów połączonych ze sobą w heterogeniczną infrastrukturę sprzętową.

Kolejnym kierunkiem badań nad poprawą efektywności obliczeń w skali makro było opracowanie metodyki redukcji złożoności obliczeniowej z wykorzystaniem

SSRVE. Opracowanie metodyki przedstawione jest szczegółowo w ramach prac [7-9]. Następnie zaproponowane zostało podejście umożliwiające połączenie SSRVE redukującego złożoność obliczeniową domeny z możliwościami architektur heterogenicznych typu HPC. Rozwiązanie to przedstawiają prace [10-12]. Dzięki temu możliwe było zastosowanie bardzo złożonych i kosztownych obliczeniowo wieloiteracyjnych procedur optymalizacji, które mogą być zastosowane w praktyce przemysłowej [13]. Ostatnim krokiem było połączenie rozwiązania SSRVE wraz z Analizą Izogeometryczną IGA, co pozwoliło na dalsze zminimalizowanie ilości elementów niezbędnych do opisu domeny obliczeniowej przy zachowaniu wysokiej jakości wyników symulacji numerycznych. Szczegółowy opis tego podejścia wraz z wynikami przedstawiają prace [14,15].

Głównym osiągnięciem jest opracowanie metodyki dekompozycji domeny obliczeniowej w zależności od dostępnej heterogenicznej architektury sprzętowej, co pozwala na taki podział zadań obliczeniowych, aby czas oczekiwania na synchronizację pomiędzy kolejnymi iteracjami był minimalny. Ponadto, znaczącym elementem całej metodyki jest zarządzanie reprezentacją domeny obliczeniowej z wykorzystaniem rozwiązań SSRVE oraz IGA. Umożliwiło to zachowanie wysokiej jakości wyników przy jednoczesnej redukcji nakładu kosztów obliczeniowych poprzez znaczące zmniejszenie ilości elementów skończonych oraz brak konieczności wprowadzania procedur remeshingu. Dodatkowo, istotne jest opracowanie podejścia do składania układów równań oraz algorytmika procedur adaptacji siatek dla heterogenicznych architektur sprzętowych.

2. Wybór optymalnego rozwiązania pomiędzy doborem architektury sprzętowej a konfiguracją wieloskalowych metod obliczeniowych w podejściu sprzężonym.

W przypadku sprzężonych wieloskalowych symulacji numerycznych wymiana danych pomiędzy skalami jest dwukierunkowa, a zatem podczas gdy jedna z metod działa w skali wyższej, wówczas metoda niższa pozostaje nieaktywna. Metody działają więc zależnie od siebie w każdym kroku czasowym, a dodatkowo w przypadku hybrydowego podejścia wieloskalowego operują na tej samej domenie obliczeniowej. W takiej sytuacji należy zwrócić szczególną uwagę na czas związany z alokacją danych na urządzeniu obliczeniowym, metody interpolacji danych, które mogą działać na tym urządzeniu oraz dynamiczne zarządzanie uruchamianiem kerneli obliczeniowych. Szczególny nacisk położony został na statyczne oraz dynamiczne równoważenie obciążenia obliczeniowego, które jest kluczowe, aby zminimalizować czas jałowe metod w poszczególnych skalach. Na bazie wyników uzyskanych dla pojedynczych skali, wskazujących dobór parametrów uruchomieniowych oraz dobór architektury sprzętowej w celu uzyskania optymalnych czasów obliczeń dla przeanalizowanych metod, opracowano rozwiązania umożliwiające podział obliczeń ze względu na dwa następujące poziomy zrównoleglenia:

- a. Pomędzy urządzeniami obliczeniowymi w ramach jednego węzła – charakterystyki obliczeniowe wyznaczone w poprzednim punkcie dla obliczeń w pojedynczych skalach pozwoliły na wskazanie jak powinny być uruchamiane obliczenia wieloskalowe jednocześnie na CPU będącym gospodarzem obliczeń jak i na urządzeniach obliczeniowych w ramach jednego węzła. Dzięki temu gospodarz będący w stanie spoczynku w trakcie trwania obliczeń na urządzeniu może być wykorzystywany do przeprowadzania symulacji numerycznych w innej skali w tym samym czasie. Sposób podziału danych i obliczeń w obrębie jednego węzła przedstawia praca [5].

Głównym osiągnięciem jest opracowanie metody, pozwalającej na określenie optymalnej wielkości zadania obliczeniowego dla hybrydowych metod wieloskalowych w pełni sprzężonych, które pracują jednocześnie na jednej domenie obliczeniowej realizując różną funkcjonalność obliczeniową w różnych skalach. Przykład takiego rozwiązania dla metody CAFE został przedstawiony szczegółowo w publikacji [16].

- b. Pomędzy węzłami obliczeniowymi – dzięki wynikom z poprzedniego punktu uzyskanym dla pojedynczych węzłów obliczeniowych możliwa była odpowiedź na pytanie, w jaki sposób maksymalnie wykorzystać zasoby komputera z wieloma procesorami i urządzeniami liczącymi. W przypadku wielu takich węzłów zastosowano algorytm do minimalizacji kosztu komunikacyjnego, aby jeszcze poprawić wydajność kompleksowego algorytmu wieloskalowego dla kilku węzłów. Wyniki tej pracy opisane zostały w artykule [6], który przedstawia podejście do statycznego i dynamicznego równoważenia obciążenia dla modeli wieloskalowych.

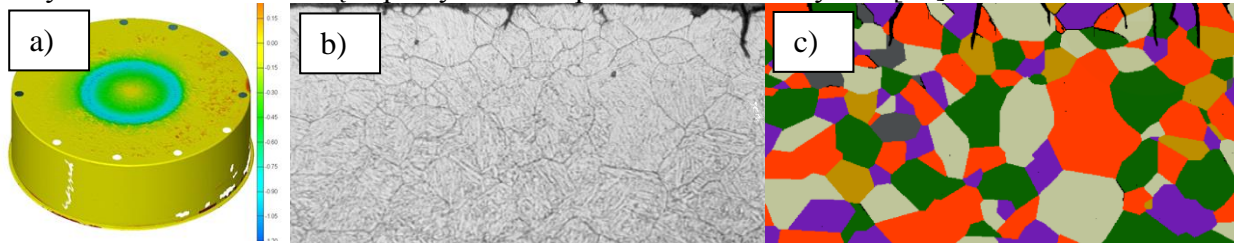
Głównym osiągnięciem jest nowa procedura dynamicznego równoważenia obciążenia obliczeniowego dedykowana obliczeniom wieloskalowym, która pozwala w znacznym stopniu zrównoważyć już pierwszą iterację obliczeń, gdzie inne procedury muszą oczekiwać na pierwszy wynik obliczeń i dopiero na ich podstawie proponują zoptymalizowane rozwiązanie. Stworzona metodyka obejmuje przede wszystkim sposób przygotowania procedur numerycznych, sposób zrównoleglenia domeny obliczeniowej, statyczne oraz dynamiczne równoważenie obciążenia obliczeniowego oraz sposób realizacji samych obliczeń wieloskalowych na heterogenicznych węzłach sprzętowych z kilkoma urządzeniami. W ramach prac zaimplementowana została metoda *Factorial Design* umożliwiająca analizę wpływu parametrów algorytmów, urządzeń i węzłów obliczeniowych na wydajność podejścia wieloskalowego. Zaproponowana metodyka opracowana została głównie dla obliczeń wykorzystujących podejście CAFE (ang. Cellular Automata Finite Element), gdzie automaty komórkowe wykorzystywane były do modelowania skali mikro natomiast metoda elementów skończonych do modelowania skali makro. W ramach tej analizy opracowane zostały prace [16,17], przedstawiające możliwość wykorzystania architektur heterogenicznych w modelowaniu wieloskalowym. Prace te były głównym przyczynkiem do opracowania artykułu [6], w którym przedstawiono zastosowanie inspirowanych naturą metod optymalizacji do dynamicznego równoważenia obciążenia obliczeniowego w architekturach złożonych ze sprzętu obliczeniowego o różnej charakterystyce wydajnościowej.

Uzyskane wyniki i wykorzystanie w praktyce

W ramach pracy przeanalizowane zostały metody MS/MD do prowadzenia symulacji numerycznych skali nano, CA w skali mikro oraz FE w skali mikro oraz makro. Dodatkowo zaimplementowane zostały moduły interfejsowe pozwalające na łączenie obliczeń w kilku skalach poprzez interpolację wartości z jednej domeny obliczeniowej do drugiej. Moduły te zostały zaimplementowane w formie kerneli obliczeniowych, dzięki czemu nie ma konieczności wysyłania danych do urządzenia gospodarza (host) z wykorzystaniem wąskiego gardła, jakim jest magistrala PCI Express. Opracowane metody wraz z interfejsami komunikacyjnymi pozwoliły na stworzenie rozwiązań wieloskalowych umożliwiających przeprowadzenie symulacji numerycznych typu nano-makro (MS-FE), nano-mikro-makro (MS-CAFE) oraz mikro-makro (CAFE i FE²). Największą popularnością wśród zastosowań przemysłowych cieszą się obecnie ostatnie z wymienionych rozwiązań, ponieważ wśród nich istnieje jeszcze największy potencjał w kierunku przyspieszania zadań obliczeniowych.

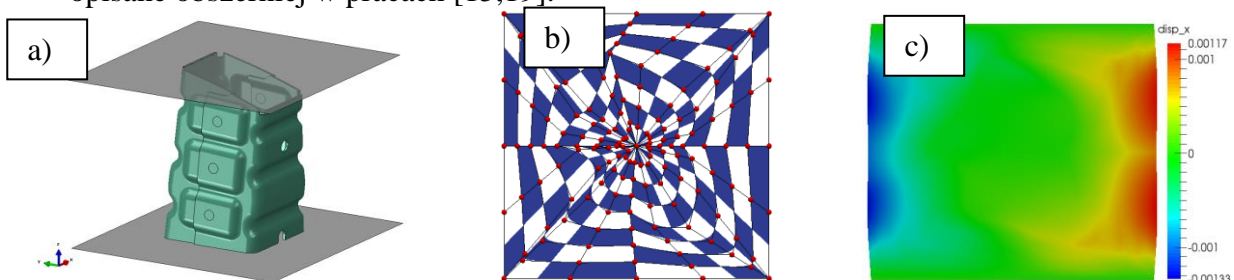
Weryfikacja opracowanego rozwiązania pozwoliła otrzymać następujące wyniki jakościowe dla różnych podejść:

- Pół-sprężone obliczenia wieloskalowe mikro-makro w zastosowaniu do symulacji pęknięcia zmęczeniowego w procesie kucia matrycowego na gorąco. W rozwiązaniu zastosowane zostało podejście wieloskalowe (skala makro – metoda FE zastosowana do modelowania naprężeń w zakresie sprężystym, skala mikro – platforma do automatów komórkowych CA z modelem pęknięcia zmęczeniowego). Wynik jakościowy przedstawia rysunek 1. Pozostała część pracy została opublikowana w artykule [18].



Rys. 1. Wynik modelowania pęknięcia zmęczeniowego w skali mikro. a) matryca w skali makro, b) oryginalna mikrostruktura, c) wynik symulacji.

- W pełni sprężone obliczenia wieloskalowe mikro-makro w zastosowaniu do symulacji tłoczenia słupków zderzeniowych (ang. *crash-box*) ze stali dwufazowej, będących elementami nadwozia samochodowego odpowiedzialnymi za pochłanianie energii w czasie wypadków (skala makro – metoda FE zastosowana do modelowania przemieszczeń, skala mikro – metoda FE oraz IGA zastosowane do modelowania przemieszczeń, naprężeń i odkształceń). Rysunek 2 przedstawia model oraz wykorzystanie metody IGA. Wyniki są opisane obszerniej w pracach [15,19].



Rys. 2. Wynik modelowania dwufazowej stali DP w skali mikro podczas deformacji elementów nadwozia samochodowego. a) model słupka zderzeniowego, b) siatka IGA na domenie obliczeniowej statystycznie podobnego reprezentatywnego elementu objętościowego w skali mikro, c) wynik symulacji – przemieszczenia w skali mikro.

- Obliczenia w pojedynczej skali zostały zweryfikowane dla statyki i dynamiki molekularnej, automatów komórkowych oraz metody elementów skończonych dając poprawne wyniki jakościowe.

Szczegółowy opis pozostałych praktycznych zastosowań opracowanej metodyki w aspekcie efektywności obliczeniowej oraz wykorzystania heterogenicznych architektur HPC w modelowaniu złożonych procesów i cykli przeróbki plastycznej metali znajduje się w artykułach [20-23]. Elementy opracowanej metodyki wraz z praktycznym opisem zastosowań oraz szczegółami uzyskanych wyników opisane są w książce [24].

Szczegóły implementacyjne

- Modele obliczeniowe w skali nano – OpenCL 2.0 oraz CUDA
- Platforma CA – C++, modele obliczeniowe w skali mikro – OpenCL 2.0

- Modele obliczeniowe w skali makro OpenCL 2.0 oraz CUDA
- Interfejsy interpolujące dane w celu połączenia kilku skali – OpenCL 2.0

Wpływ na dyscyplinę naukową

Opracowanie tytułowej metodyki obliczeń stworzyło podstawy do rozwoju dyscypliny **informatyka** w następujących kierunkach:

- Algorytmika metod do wieloskalowych symulacji numerycznych – przeniesienie opracowanego podejścia na grunt innych metod numerycznych stosowanych w sposób sprzężony lub w pół-sprzężony umożliwi poprawę ich efektywności obliczeniowej, a tym samym optymalizację zasobów na węzłach obliczeniowych w dużych infrastrukturach sprzętowych. Ponadto, wciąż otwarty jest rozwój symulacji numerycznych na bazie analizy izogeometrycznej, który może znaleźć zastosowanie w niekonwencjonalnych architekturach sprzętowych.
- Algorytmika metod wspomagających obliczenia wieloskalowe – rozwój metod adaptacji lub interpolacji domen obliczeniowych oraz analiza ich wpływu na ostateczną efektywność obliczeniową stanowi jeden z elementów wpływających na minimalizację czasu oczekiwania na wyniki obliczeń w poszczególnych skalach.
- Zastosowania nowoczesnych energooszczędnych architektur sprzętowych z jednostkami typu ARM lub FPGA w modelowaniu wieloskalowym – kierunek ten jest zgodny z obecnymi trendami Green Computing, a opracowana metodyka stanowi podstawę do stworzenia ogólnego podejścia umożliwiającego dynamiczne zarządzanie konfiguracją nowoczesnych architektur sprzętowych i algorytmów w celu minimalizacji zużycia energii elektrycznej niezbędnej do przeprowadzenia obliczeń.
- Komponentowe symulacje numeryczne – kierunek ten miałby na celu opracowanie rozwiązania pozwalającego na wykorzystanie gotowych homogenicznie implementowanych modułów obliczeniowych dla heterogenicznych architektur sprzętowych w celu łatwej konfiguracji i uruchomienia złożonych obliczeń wieloskalowych. Pozwoliłoby to na łatwiejszą aplikowalność i optymalizację takich modułów w nowoczesnych e-infrastrukturach. Opracowane metody stanowią dobrą podstawę do rozpoczęcia takich badań.

Podsumowanie

W ramach badań naukowych opracowana została metodyka umożliwiająca:

- a. Wyznaczenie optymalnej konfiguracji metod obliczeniowych (np.: podziału domeny obliczeniowej, reprezentacji danych, doboru procedur numerycznych obliczeniowych i wspomagających, synchronizacji obliczeń) wówczas, gdy znana jest architektura sprzętu, z pomocą którego wykonane zostaną obliczenia.
- b. Określenie optymalnego sprzętu komputerowego spośród dostępnej puli dla określonych symulacji numerycznych.

Obecnie nie ma rozwiązań w literaturze naukowej, które kompleksowo pozwalałyby na analizę złożonych metod wieloskalowych dla różnych urządzeń obliczeniowych. Przedstawiona w niniejszym dokumencie metodyka pozwala na znaczne skrócenie czasu złożonych obliczeń wieloskalowych, zarówno w podejściu hierarchicznym jak i hybrydowym, które mogą być sprzężone lub półsprzężone. Poprawa efektywności następuje przede wszystkim poprzez minimalizację czasów jałowych, które mają miejsce przed synchronizacją danych pomiędzy poszczególnymi skalami. Dzięki temu metodyka wpływa nie tylko na czas oczekiwania na obliczenia, ale również na znaczną poprawę energooszczędności. Jest szczególnie widoczne podczas obliczeń z wykorzystaniem kart graficznych i kart

obliczeniowych typu Xeon Phi, ale także nowoczesnych procesorów ARM w połączeniu z mobilnymi kartami graficznymi oraz procesorów FPGA. Jednym z istotniejszych aspektów jest fakt homogenicznej implementacji poszczególnych modułów obliczeniowych oraz uniwersalność opracowanego podejścia, dzięki czemu oferuje ono szeroki wachlarz przyszłych kierunków rozwoju naukowego na gruncie informatyki.

Na podstawie wyników badań związanych z tytułową metodyką opracowane zostały trzy nowe wnioski projektowe:

- a. Wniosek do konkursu CHIST-ERA pt. *DYNamic Adaptation for Multiscale numerical simulations on heterogeneous distributed cOmputing platforms* złożony we współpracy z prof. Hong-Linh Truong (Vienna University of Technology) oraz prof. Sven-Bodo Scholz (Heriot Watt University).
- b. Wniosek na konkurs Narodowego Centrum Nauki Sonata Bis pt. *Green Multiscale Computing – energooszczędna strategia obliczeniowa dla wieloskalowych symulacji numerycznych*.
- c. Wniosek na konkurs Narodowego Centrum Nauki OPUS pt. *Zastosowanie systemów opartych o wiedzę do kontrolowania niepewności w optymalizacji procesów przetwórstwa metali*.

5. Omówienie pozostałych osiągnięć naukowo-badawczych.

W ramach mojej pracy naukowej zajmowałem się wieloma aspektami informatyki oraz jej zastosowania w różnych gałęziach nauki. Po ukończeniu obronie pracy doktorskiej dotyczącej algorytmów przetwarzania i filtrowania danych wielowymiarowych zająłem się analizą obrazów, która była naturalnym rozwinięciem tematyki doktoratu w zakresie danych dwuwymiarowych. W analizie obrazów stosowałem do przetwarzania wstępnego i filtrowania swój algorytm dynamiki cząstek. Z racji miejsca zatrudnienia wykorzystywałem go przede wszystkim do obrazów mikrostruktur materiałowych, a w szczególności stali jedno- i wielofazowych oraz stopów metali jak np.: stopy magnezu stosowane do produkcji nici chirurgicznych. W tej dziedzinie udało mi się opracować i opublikować kilka algorytmów do segmentacji i klasteryzacji obrazów. Dzięki tym rozwiązaniom możliwe było otrzymanie wirtualnej reprezentacji materiału bezpośrednio ze zdjęcia mikroskopowego, co pozwoliło znacznie zautomatyzować pracę i rozszerzyć reprezentację materiału nie tylko o wygląd ale także o własności. Tak kompleksowe podejście do opisu materiału nazywa się Cyfrową Reprezentacją Materiału (ang. Digital Material Representation, DMR).

W badaniach związanych z DRM uczestniczyłem aktywnie przy tworzeniu modeli numerycznych bazujących na cyfrowej reprezentacji materiału oraz wizualizacji DMR. W ramach tych badań powstało kilkanaście publikacji w czasopismach i na konferencjach. Zaowocowało to powstaniem rozwiązania do wizualizacji DMR dla projektorów 3D oraz opracowaniem algorytmu do obsługi bardzo dużych struktur materiałowych zapisanych w plikach. Wspomniany algorytm umożliwiał inteligentne buforowanie danych z dysku, dzięki czemu możliwa była płynna interaktywna wizualizacja materiałów opisanych w różnych skalach. Ponadto, opracowałem koncepcję platformy programistycznej umożliwiającej tworzenie modeli materiałowych DMR z wykorzystaniem predefiniowanych struktur programistycznych. Opracowany framework stał się podstawą do opracowania wielu publikacji obejmujących różne modele materiałowe oraz aspekty efektywności obliczeniowej jak zrównoleglenie i rozproszenie obliczeń, a także zastosowanie idei *workflow*.

Rozwiązania wypracowane w tematyce DMR oraz wizualizacji zapoczątkowały moją pracę z modelowaniem wieloskalowym, w ramach której opracowałem i opublikowałem kilka modeli numerycznych oraz przedstawiłem możliwości ich zastosowania w praktyce przemysłowej. Były to modele zrealizowane w oparciu o konwencjonalne podejście wieloskalowe jak i również w oparciu o DMR, zarówno dla wieloskalowego podejścia hierarchicznego jak i hybrydowego. Problematyka obliczeniowa związana z opracowanymi przeze mnie rozwiązaniami stała się podstawą do dalszej analizy od strony efektywności numerycznej oraz możliwości zastosowania niekonwencjonalnych architektur sprzętowych. Wyniki tej analizy są obecnie głównym osiągnięciem naukowym przedstawionym w niniejszym autoreferacie.

Niezależnie od wspomnianych powyżej zainteresowań związanych z DMR oraz modelowaniem wieloskalowym zaangażowałem się w projektowanie oraz implementację systemów komputerowych wspomagających zarządzanie przedsiębiorstwami oraz projektowanie procesów przemysłowych. Zainteresowania te są zbieżne z drugim kierunkiem studiów, który ukończyłem tj. z zarządzaniem. Jestem współautorem rozwiązania typu CRM (*Customer Relationship Management*), który został sprzedany i wdrożony w kilkunastu przedsiębiorstwach w Polsce. W pracy naukowej zająłem się jednak aspektem wspomagania projektowania produkcji. Systemy, które opracowywałem, wykorzystują nie tylko modelowanie i symulacje numeryczne, ale przede wszystkim optymalizację i analizę wrażliwości, które mają największe znaczenie w projektowaniu parametrów procesów i cykli przemysłowych. Jednak kluczowy z punktu widzenia przemysłu okazał się czas wykonywania obliczeń i konieczność długiego oczekiwania na precyzyjne wyniki. Dlatego w kolejnych

moich pracach skupiłem się na poprawie efektywności obliczeniowej procedur numerycznych oraz wykorzystaniu infrastruktury HPC w optymalizacji i analizie wrażliwości. W tym zakresie badań powstało wiele projektów, publikacji naukowych oraz wdrożeń. Oprócz wymienionych rozwiązań naukowych ciekawym osiągnięciem było opracowanie systemu komputerowego ManuOpti, który pełni rolę integratora zewnętrznych programów do symulacji numerycznych oraz bibliotek numerycznych do analizy wrażliwości i optymalizacji. System jest rozwiązaniem praktycznym wciąż udoskonalanym o kolejne funkcjonalności. Jedną z nich jest kierunek związany z metamodelowaniem, przy którym uczestniczyłem w tworzeniu zarówno oprogramowania do zarządzania metamodelami jak i w tworzeniu samych metamodeli zbudowanych w oparciu o sieci neuronowe. Wynikiem tych prac jest kilka publikacji w czasopiśmie oraz wystąpienia na międzynarodowych i krajowych konferencjach naukowych.